

The SEQ ID NO:7 primer contains sequences encoding a 16-amino acid Gly-Ser peptide linker including a *Bam*HI restriction enzyme site. The resulting DNA fragments of approximately 700 bp in length are inserted into a holding vector such as pUC19 at the *Bam*HI and *Eco*RI sites to give the pL-vFc γ 2 plasmid. The sequence of the gene is confirmed by DNA sequencing.

[0018] To prepare the HuEPO-L-vFc γ 2 fusion gene, the EPO fragment is excised from the pEPO plasmid with *Hind*III and *Bam*HI and is purified by agarose gel electrophoresis. The purified fragment is then inserted to the 5'-end of the peptide linker in the pL-vFc γ 2 plasmid to give the pEPO-L-vFc γ 2 plasmid. The fusion gene comprises HuEPO, a Gly-Ser peptide linker and the Fc γ 2 variant gene.

[0019] The presence of a peptide linker between the EPO and Fc moieties increases the flexibility of the EPO domains and enhances its biological activity (see, for example, Sytkowski et al., *J. Biol. Chem.*, 274: 24773-8, 1999). For the present invention, a peptide linker of about 20 or fewer amino acids in length is preferred. Peptide linker comprising two or more of the following amino acids: glycine, serine, alanine, and threonine can be used. An example of the peptide linker contains Gly-Ser peptide building blocks, such as GlyGlyGlyGlySer. Figure 2A shows a fusion gene (SEQ ID NO. 17) containing sequences encoding HuEPO, a 16-amino acid peptide linker (GlySerGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer, SEQ ID NO: 23), and the Fc γ 2 Pro331Ser variant, and its corresponding amino acid sequence (SEQ ID NO. 18).

[0020] The complete gene encoding the HuEPO-L-vFc fusion protein is then inserted at the *Hind*III and *Eco*RI sites of a mammalian expression vector, such as pcDNA3 (Invitrogen). The final expression vector plasmid, named pEFP2, contains the cytomegalovirus early gene promoter-enhancer which is required for high level expression in mammalian cells. The plasmid also contains selectable markers to confer ampicillin resistance in bacteria, and G418 resistance in mammalian cells. In addition, the pEFP2 expression vector contains the dihydrofolate reductase (DHFR) gene to enable the co-amplification of the HuEPO-L-vFc γ 2 fusion gene and the DHFR gene in the presence of methotrexate (MTX)

[0023] To prepare the HuEPO-L-vFc_{γ4} fusion gene, the HuEPO fragment is excised from the pEPO plasmid with *Hind*III and *Bam*HI and then inserted to the 5'-end of the peptide linker in the pL-vFc_{γ4} plasmid to give the pEPO-L-vFc_{γ4} plasmid. This fusion gene comprising HuEPO, a 16-amino acid Gly-Ser peptide linker and the Fc_{γ4} variant gene is then inserted at the *Hind*III and *Eco*RI sites of a mammalian expression vector, such as pcDNA3 (Invitrogen), as described for the HuEPO-L-vFc_{γ2} fusion protein. The final expression vector plasmid is designated as pEFP4. Figure 2B shows a fusion gene (SEQ ID NO. 19) containing sequences encoding HuEPO, a 16-amino acid peptide linker (GlySerGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer, SEQ ID NO: 23), and the Fc_{γ4} variant with Ser228Pro and Leu235Ala mutations, and its corresponding amino acid sequence (SEQ ID NO. 20).

1. Construction of the gene encoding the HuEPO-L-vFc_{γ1} fusion protein

C²
[0024] The hinge domain of human IgG1 heavy chain contains 15 amino acid residues (GluProLysSerCysAspLysThrHisThrCysProProCysPro, SEQ ID NO: 24) including 3 cysteine residues. Out of these 3 cysteine residues, the 2nd and 3rd are involved in the formation of disulfide bonding between two heavy chains. The 1st cysteine residue is involved in the disulfide bonding to the light chain of IgG. Since there is no light chain present in the Fc fusion protein molecule, this cysteine residue may pair with other cysteine residues, leading to nonspecific disulfide bonding. The hinge domain of Fc_{γ1} can be truncated to eliminate the 1st cysteine residue (AspLysThrHisThrCysProProCysPro, SEQ ID NO: 25). The gene encoding the Fc_{γ1} region is obtained by reverse transcription and PCR using RNA prepared from human leukocytes and appropriate 5' primer (SEQ ID NO:13) and 3' primer (SEQ ID NO:4). Resulting DNA fragments containing the truncated hinge and complete sequences of CH2 and CH3 domains of Fc_{γ1} is used as the template to generate the Fc_{γ1} variant with Leu234Val, Leu235Ala, and Pro331Ser mutations (vFc_{γ1}).

[0025] One way to incorporate these mutations is as follows: two segments are produced and then assembled by using the natural Fc_{γ1} as the template in overlapping PCR. The 5' segment is generated by using SEQ ID NO:14 as the 5'

11

primer and SEQ ID NO:5 as the 3' primer. This 5' primer contains the Leu234Val, Leu235Ala mutations and the 3' primer contains the Pro331Ser mutation. The 3' segment is generated by using SEQ ID NO:6 as the 5' primer and SEQ ID NO:4 as the 3' primer. These 5' and 3' segments are then joined at the region covering the Pro331Ser mutation by using SEQ ID NO:14 as the 5' primer and SEQ ID NO:4 as the 3' primer. This amplified fragment of approximately 650 bp in length, together with a synthetic oligonucleotide of 55 bases (SEQ ID NO:15) containing Leu234Val and Leu235Ala, are joined in PCR by using SEQ ID NO:16 as the 5' primer and SEQ ID NO:4 as the 3' primer. The SEQ ID NO:16 primer contains sequences encoding a 16-amino acid Gly-Ser peptide linker including the *Bam*HI site. The resulting DNA fragments of approximately 700 bp in length are inserted into a holding vector such as pUC19 at the *Bam*HI and *Eco*RI sites to give the pL-vFc_γ1 plasmid. The sequence of the gene is confirmed by DNA sequencing.

15 [0026] To prepare the HuEPO-L-vFc_γ1 fusion gene, the EPO fragment is excised from the pEPO plasmid with *Hind*III and *Bam*HI and inserted to the 5'-end of the peptide linker in the pL-vFc_γ1 plasmid to give the pEPO-L-vFc_γ1 plasmid. The fusion gene comprising HuEPO, a 16-amino acid Gly-Ser peptide linker, and the Fc_γ1 variant gene is then inserted at the *Hind*III and *Eco*RI sites of a mammalian expression vector, such as pcDNA3 (Invitrogen), as described for the HuEPO-L-vFc_γ2 fusion protein. The final expression vector plasmid is designated as pEFP1. Figure 2C shows a fusion gene (SEQ ID NO. 21) containing sequences encoding HuEPO, a 16-amino acid peptide linker (GlySerGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer, SEQ ID NO: 23), and the Fc_γ1 variant with Leu234Val, Leu235Ala and Pro331Ser mutations, and its corresponding amino acid sequence (SEQ ID NO. 22).

2. Expression of the fusion protein in transfected cell lines

30 [0027] The recombinant pEFP1, pEFP2 or pEFP4 expression vector plasmid is transfected into a mammalian host cell line to achieve the expression of the HuEPO-L-vFc fusion protein. For stable high levels of expression, a preferred host cell line is Chinese Hamster Ovary (CHO) cells deficient in the DHFR enzyme (see, for example, US patent 4,818,679). A preferred method of transfection is

Figure 1. Amino acid sequence alignment in human IgG isotypes and their variants.

| Human IgG Isotype | Amino Acid Position | | | | | |
|-------------------|---------------------|------------|-------|-------------|------------|--|
| | 228.....234 | 235 | 236 | 237.....330 | 331 | |
| G1 | Pro.....Leu | Leu | Gly | Gly.....Ala | Pro | |
| G2 | Pro.....Val | Ala | | Gly.....Ala | Pro | |
| G4 | Ser.....Phe | Leu | Gly | Gly.....Ser | Ser | |
| G1 variant | Pro..... Val | Ala | Gly | Gly.....Ala | Ser | |
| G2 variant | Pro.....Val | Ala | | Gly.....Ala | Ser | |
| G4 variant | ProPhe | Ala | Gly | Gly.....Ser | Ser | |

ID number

Corresponding Row in this Figure 1

SEQ ID NO:26

G1

SEQ ID NO:27

G2

SEQ ID NO:28

G4

SEQ ID NO:22

G1 variant

SEQ ID NO:18

G2 variant

SEQ ID NO:20

G4 variant

Figure 2A. DNA and deduced amino acid sequences of HuEP0-L-vFC_{γ2}

| DNA | SEQ ID NO: 17 |
|---------------------|---------------|
| Amino Acid Sequence | SEQ ID NO: 18 |

```

aag ctt ggc gcg gag atg ggg gtg cac gaa tgt cct gcc tgg ctg ttg ctt ctg ctg tcc      60
HindIII      M G V H E C P A W L W L L S
-27
ctg ctg tcg ctc cct ctg ggc ctc cca gtc ctg ggc gcc cca cca cgc ctc atc tgt gac     120
L L S L P L G L P V L G A P P R L I C D
-10
agc cga gtc ctg gag agg agg tac ctc ttg gag gcc aag gag gcc gag aat atc acg acg ggc     180
S R V L E R Y L L E A K E A E N I T T G
10
tgt gct gaa cac tgc agc ttg aat gag aat atc act gtc cca gac acc aaa gtt aat ttc     240
C A E H C S L N E N I T V P D T K V N F
30
tat gcc tgg aag agg atg gag gtc ggg cag cag gcc gta gaa gtc tgg cag ggc ctg gcc     300
Y A W K R M E V G Q Q A V E V W Q G L A
50
ctg ctg tcg gaa gct gtc ctg cgg ggc cag gcc ctg ttg gtc aac tct tcc cag ccg tgg     360
L L S E A V L R G Q A L L V N S S Q P W
70
gag ccc ctg cag ctg cat gtg gat aaa gcc gtc agt ggc ctt cgc agc ctc acc act ctg     420
E P L Q L H V D K A V S G L R S L T T L
90
ctt cgg gct ctg gga gcc cag aag gaa gcc atc tcc cct cca gat gcg gcc tca gct gct     480
L R A L G A Q K E A I S P P D A A S A A
110
cca ctc cga aca atc act gct gac act ttc cgc aaa ctc ttc cga gtc tac tcc aat ttc     540
P L R R T I T A D T F R K L F R V Y S N F
130
ctc cgg gga aag ctg aag ctg tac aca ggg gag gcc tgc agg aca ggg gac gga tcc ggt     600
L R G K L K L Y T G E A C R T G D G S G
150
ggc ggt tcc ggt gga ggc gga agc ggc ggt gga gga tca gag cgc aaa tgt tgt gtc gag     660
G G S G G G S G G G S S E R K C C V E
170

```

09/932.812

tgc cca ccg tgc cca gca cca cct gtg gca gga ccg tca gtc ttc ctc ttc ccc cca aaa 720
 C P P C P A P P V A G P S V F L F P P K
 190
 ccc aag gac acc ctc atg atc tcc cgg acc cct gag gtc acg tgc gtg gtg gtg gac gtg 780
 P K D T L M I S R T P E V T C V V V D V
 210
 agc cac gaa gac ccc gag gtc cag ttc aac tgg tac gtg gac ggc gtg gag gtg cat aat 840
 S H E D P E V Q F N W Y V D G V E V H N
 230
 gcc aag aca aag cca cgg gag gag cag ttc aac agc acg ttc cgt gtg gtc agc gtc ctc 900
 A K T K P R E E Q F N S T F R V V S V L
 250
 acc gtt gtg cac cag gac tgg ctg aac ggc aag gag tac aag tgc aag gtc tcc aac aaa 960
 T V V H Q D W L N G K E Y K C K V S N K
 270
 ggc ctc cca gcc tcc atc gag aaa acc atc tcc aaa acc aaa ggg cag ccc cga gaa cca 1020
 G L P A S I E K T I S K T K G Q P R E P
 290
 cag gtg tac acc ctg ccc cca tcc cgg gag gag atg acc aag aac cag gtc agc ctg acc 1080
 Q V Y T L P P S R E E M T K N Q V S L T
 310
 tgc ctg gtc aaa ggc ttc tac ccc agc gac atc gcc gtg gag tgg gag agc aat ggg cag 1140
 C L V K G F Y P S D I A V E W E S N G Q
 330
 ccg gag aac aac tac aag acc aca cct ccc atg ctg gac tcc gac ggc tcc ttc ttc ctc 1200
 P E N N Y K T T P P M L D S D G S F F L
 350
 tac agc aag ctc acc gtg gac aag agc agg tgg cag cag ggg aac gtc ttc tca tgc tcc 1260
 Y S K L T V D K S R W Q Q G N V F S C S
 370
 gtg atg cat gag gct ctg cac aac cac tac tac acg cag aag agc ctc tcc ctg tct ccg ggt 1320
 V M H E A L H N H Y T Q K S L S L S P G
 390
 aaa tga gaa ttc
 K EcoRI
 409
 1332

Figure 2B. DNA and deduced amino acid sequences of HuEPO-L-vFc_γ4

| DNA | SEQ ID NO: 19 |
|---------------------|---------------|
| Amino Acid Sequence | SEQ ID NO: 20 |

| | |
|---|-----|
| aag ctt ggc gcg gag atg ggg gtg cac gaa tgt cct gcc tgg ctg tgg ctt ctc ctg tcc | 60 |
| <i>HindIII</i> M G V H E C P A W L W L L S | |
| ctg ctg tcg ctc cct ctg ggc ctc cca gtc ctg ggc gcc cca cca cgc ctc atc tgt gac | 120 |
| L L S L P L G L P V L G A P P R L I C D | |
| -10 | |
| agc cga gtc ctg gag agg tac ctc ttg gag gcc aag gag gcc gag aat atc acg acg ggc | 180 |
| S R V L E R Y L L E A K E A E N I T T G | |
| 10 | |
| tgt gct gaa cac tgc agc ttg aat gag aat atc act gtc cca gac acc aaa gtt aat ttc | 240 |
| C A E H C S L N E N I T V P D T K V N F | |
| 30 | |
| tat gcc tgg aag agg atg gag gtc ggg cag cag gcc gta gaa gtc tgg cag ggc ctg gcc | 300 |
| Y A W K R M E V G Q Q A V E V W Q G L A | |
| 50 | |
| ctg ctg tcg gaa gct gtc ctg cgg ggc cag gcc ctg ttg gtc aac tct tcc cag ccg tgg | 360 |
| L L S E A V L R G Q A L L V N S S Q P W | |
| 70 | |
| gag ccc ctg cag ctg cat gtg gat aaa gcc gtc agt ggc ctt cgc agc ctc acc act ctg | 420 |
| E P L Q L H V D K A V S G L R S L T T L | |
| 90 | |
| ctt cgg gct ctg gga gcc cag aag gaa gcc atc tcc cct cca gat gcg gcc tca gct gct | 480 |
| L R A L G A Q K E A I S P P D A A S A A | |
| 110 | |
| cca ctc cga aca atc act gct gac act ttc cgc aaa ctc ttc cga gtc tac tcc aat ttc | 540 |
| P L R T I T A D T F R K L F R V Y S N F | |
| 130 | |
| ctc cgg gga aag ctg aag ctg tac aca ggg gag gcc tgc agg aca ggg gac gga tcc ggt | 600 |
| L R G K L K L Y T G E A C R T G D G S G | |
| 150 | |
| ggc ggt tcc ggt gga ggc gga agc ggc ggt gga gga tca gag tcc aaa tat ggt ccc cca | 660 |
| G G S G G G S G G G G G S S E S K Y G P P | |
| 170 | |

09/932.812

tgc cca gcca tgc cca gca cct gag ttc gcg ggg gga cca tca gtc ttc ctg ttc ccc cca 720
 C P P C P A P E F A G G P S V F L F P P
 190
 aaa ccc aag gac act ctc atg atc tcc cgg acc cct gag gtc acg tgc gtg gtg gtg gac 780
 K P K D T L M I S R T P E V T C V V V D
 210
 gtg agc cag gaa gac ccc gag gtc cag ttc aac tgg tac gtg gat ggc gtg gag gtg cat 840
 V S Q E D P E V Q F N W Y V D G V E V H
 230
 aat gcc aag aca aag ccg cgg gag gag cag ttc aac agc acg tac cgt gtg gtc agc gtc 900
 N A K T K P R E E Q F N S T Y R V V S V
 250
 ctc acc gtc ctg cac cag gac tgg ctg aac ggc aag gag tac aag tgc aag gtc tcc aac 960
 L T V L H Q D W L N G K E Y K C K V S N
 270
 aaa ggc ctc ccg tcc atc gag aaa acc atc tcc aaa gcc aaa ggg cag ccc cga gag 1020
 K G L P S S I E K T I S K A K G Q P R E
 290
 cca cag gtg tac acc ctg ccc cca tcc cag gag gag atg acc aag aac cag gtc agc ctg 1080
 P Q V Y T L P P S Q E E M T K N Q V S L
 310
 acc tgc ctg gtc aaa ggc ttc tac ccc agc agc gac atc gcc gtg gag tgg gag agc aat ggg 1140
 T C L V K G F Y P S D I A V E W E S N G
 330
 cag ccg gag aac aac tac aag acc acc cct ccc gtc ctg gac tcc gac ggc tcc ttc ttc 1200
 Q P E N N Y K T T P P V L D S D G S F F
 350
 ctc tac agc agg cta acc gtg gac aag agc agg tgg cag gag ggg aat gtc ttc tca tgc 1260
 L Y S R L T V D K S R W Q E G N V F S C
 370
 tcc gtg atg cat gag gct ctg cac aac cac tac aca cag aag aag agc ctc tcc ctg tct ctg 1320
 S V M H E A L H N H Y T Q K S L S L S L
 390
 ggt aaa tga gaa ttc 1335
 G K *EcoRI*
 410

Figure 2C. DNA and deduced amino acid sequences of HuEPO-L-vFc_{γ1}

| DNA | SEQ ID NO: 21 |
|---------------------|---------------|
| Amino Acid Sequence | SEQ ID NO: 22 |

| | |
|---|-----|
| aag ctt_ggc gcc gag atg ggg gtg cac gaa tgt cct gcc tgg ctg tgg ctt ctc ctg tcc | 60 |
| <i>HindIII</i> M G V H E C P A W L W L L S | |
| ctg ctg tcg ctc cct ctg ggc ctc cca gtc ctg ggc gcc cca cca cgc ctc atc tgt gac | 120 |
| L L S L P L G L P V L G A P P R L I C D | |
| -10 | |
| agc cga gtc ctg gag agg tac ctc ttg gag gcc aag gag gcc gag aat atc acg acg ggc | 180 |
| S R V L E R Y L L E A K E A E N I T T G | |
| 10 | 20 |
| tgt gct gaa cac tgc agc ttg aat gag aat atc act gtc cca gac acc aaa gtt aat ttc | 240 |
| C A E H C S L N E N I T V P D T K V N F | |
| 30 | 40 |
| tat gcc tgg aag agg atg gag gtc ggc cag gcc gta gaa gtc tgg cag ggc ctg gcc | 300 |
| Y A W K R M E V G Q Q A V E V W Q G L A | |
| 50 | 60 |
| ctg ctg tcg gaa gct gtc ctg cgg ggc cag gcc ctg ttg gtc aac tct tcc cag ccg tgg | 360 |
| L L S E A V L R G Q A L L V N S S Q P W | |
| 70 | 80 |
| gag ccc ctg cag ctg cat gtg gat aaa gcc gtc agt ggc ctt cgc agc ctc acc act ctg | 420 |
| E P L Q L H V D K A V S G L R S L T T L | |
| 90 | 100 |
| ctt cgg gct ctg gga gcc cag aag gaa gcc atc tcc cct cca gat gcg gcc tca gct gct | 480 |
| L R A L G A Q K E A I S P P D A A S A A | |
| 110 | 120 |
| cca ctc cga aca atc act gct gac act ttc cgc aaa ctc ttc cga gtc tac tcc aat ttc | 540 |
| P L R T I T A D T F R K L F R V Y S N F | |
| 130 | 140 |
| ctc cgg gga aag ctg aag ctg tac aca ggg gag gcc tgc agg aca ggg gac gga tcc ggt | 600 |
| L R G K L K L Y T G E A C R T G D G S G | |
| 150 | 160 |
| ggc ggt tcc ggt gga ggc gga agc ggc ggt gga gga tca gac aaa act cac aca tgc cca | 660 |
| G G S S G G S S G G G G S S D K T H T C P | |
| 170 | 180 |

09/932.812

ccg tgc cca gca cct gaa gtc gcc ggg gga ccg tca gtc ttc ctc ttc ccc cca aaa ccc 720
 P C P A P E V A G G P S V F L F P P K P
 190
 aag gac acc ctc atg atc tcc cgg aca cct gag gtc aca tgc gtg gtg gtg gac gtg agc 780
 K D T L M I S R T P E V T C V V V D V S
 210
 cac gaa gac cct gag gtc aag ttc aac tgg tac gtg gac ggc gtg gag gtg cat aat gcc 840
 H E D P E V K F N W Y V D G V E V H N A
 230
 aag aca aag ccg cgg gag gag cag tac aac agc acg tac cgg gtg gtc agc gtc ctc acc 900
 K T K P P R E E Q Y N S T Y R V V S V L T
 250
 gtc ctg cac cag gac tgg ctg atc ggc aag gag tac aag gag tcc aag tgc aag gtc tcc aac aaa gcc 960
 V L H Q D W L N G K E Y K C K V S N K A
 270
 ctc cca gcc tcc atc gag aaa acc atc tcc aaa gcc aaa ggg cag ccc cga gaa cca cag 1020
 L P A S I E K T I S K A K G Q P R E P Q
 290
 gtg tac acc ctg ccc cca tcc cgg gat gag ctg acc aag aac cag gtc agc agc ctg acc tgc 1080
 V Y T L P P S R D E L T K N Q V S L T C
 310
 ctg gtc aaa ggc ttc tat ccc agc gac atc gcc gtg gag tgg gag agc aat ggg cag ccg 1140
 L V K G F Y P S D I A V E W E S N G Q P
 330
 gag aac aac tac aag acc acg cct ccc gtg ctg gac tcc gac ggc tcc ttc ttc ctc tac 1200
 E N N Y K T T P P V L D S D G S F F L Y
 350
 agc aag ctc acc gtg gac aag agc agg tgg cag ggg aac gtc ttc tca tgc tcc gtg 1260
 S K L T V D K S R W Q Q G N V F S C S V
 370
 atg cat gag gct ctg cac aac cac tac acg cag aag agc ctc tcc ctg tct ccg ggt aaa 1320
 M H E A L H N H Y T Q K S L S L S P G K
 390
 tga gaa ttc
 EcoRI
 400
 1329